Species Standalone Custom and Advanced Custom Panel Design for In Situ Xenium v1 Panels

Introduction

Xenium In Situ enables researchers to characterize RNAs in cells and tissues with subcellular resolution. Users can choose among several pre-designed panels, or design add-on custom or standalone custom panels for human and mouse. This Technical Note explains standalone custom panel design for species beyond those currently offered by 10x Genomics ("species standalone custom") and advanced custom designs for Xenium v1 chemistry. Potential advanced custom applications include isoforms, gene fusions, viral or bacterial sequences, protein tags, fluorescent reporters, and transgenes. Types of customization possible, the basic structure of Xenium probes, and how users can select their own RNA-binding domain sequences are covered here. To submit an advanced custom design request, researchers must submit a CSV file, described below, to 10x Genomics. Before creating this file, users are advised to consider these guidelines for probe design. Although pre-designed panels have been extensively tested by 10x, species standalone custom and advanced customization are not fully supported — 10x Genomics will support adding these probes to the panel, but cannot guarantee that these probes will successfully detect their targets.

Species Standalone Custom

Species standalone custom panels are panels that target endogenous gene expression in species beyond those currently offered by 10x Genomics, or panels that include multiple species. If you are interested in a non-diploid species, contact your Sales Executive.

Diploid Eukaryotic Species

To profile endogenous gene expression in diploid eukaryotic species, it is required to have a high-quality genome assembly and annotations with a well-matched single cell reference. 10x Genomics does not offer supported sample preparation protocols for species besides human and mouse. 10x expects human and mouse protocols to be adaptable to corresponding tissue types for other mammals, particularly rodents and primates, but this is not guaranteed. Researchers working with increasingly divergent lineages from mammals, particularly fish, insects, and plants, should carefully consider the impact of their unsupported sample preparation protocols on overall experimental success.

Researchers will also need to consider downstream analysis needs. For example, Xenium cell segmentation algorithms have not been tested, and are not supported, on datasets beyond specific human and mouse tissues. Researchers should be prepared to segment cells manually or with third-party tools.

Xenografts

Xenografts are transplants of organs, tissues, or cells to another species, most commonly human and mouse. 10x Genomics has built a set of gene expression designs that take human and mouse sequences into account. In general, 10x expects that multi-species experiments should work, but have not been tested at this time.

Types of Advanced Customization

Isoforms / Gene Fusions

Isoform or gene fusion designs are similar in that they both target a specific junction of sequences. Preliminary data suggest that isoform detection with Xenium is robust. Probes can be designed to target each splice junction of interest following the guidelines laid out in this Technical Note. A general recommendation is to look for an optimal ligation junction +/- 5 bp from the splice junction (see ligation junction, below), and then choose a probe that binds 20 bp upstream/downstream from that optimal junction. If no optimal junctions exist within 5 bp, neutral junctions can be tried next. Analysis of the mouse and human transcriptomes suggest that >99.9% of splice junctions have one of the neutral or preferred ligation junctions within five bases of the splice junction.

Protein Tags, Fluorescent Reporters & Transgenes

All types of exogenous sequences can be targeted with Xenium. When targeting exogenous sequences, follow the design guidelines laid out in this Technical Note. Some sequences, such as protein tags, may be shorter than 40 bp. In those cases, it is possible to shorten the oligo as long as the melting temperature stays within range. Another important consideration when designing against exogenous sequences is that these sequences are often highly expressed, which can cause optical crowding. Using the flag in the CSV to report the expression level can be important here.

Bacterial & Viral Sequences

Detection of bacterial and viral sequences has not been tested by 10x Genomics. The biggest predicted risk is that the expression level may be very high, causing optical crowding. While designing probes for bacterial sequences is possible, sample preparation for these experiments has not been tested. RNA accessibility in bacteria can be biased by the sample preparation protocol, with some bacterial species being especially difficult to lyse.

Probe Design Guidelines

To continue the design process, 10x Genomics uses an internal panel design tool (the **panel designer**). When developing an advanced customization experiment, there are specific guidelines that must be considered for designing probes.

Probe Structure

Xenium v1 padlock probes consist of two RNA binding domains connected by the padlock backbone. Upon a successful annealing event, ligation of the two RNA binding domains occurs. Each RNA binding arm of the padlock is 20 bases in length, therefore 40 bases of mRNA homology is required for successful probe design.

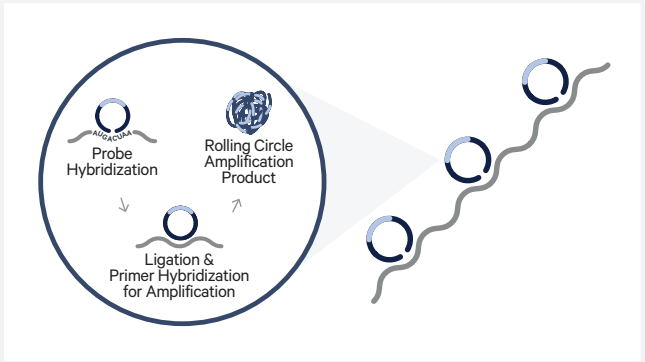

Figure 1. An individual probe has 20 bp of transcript homology on either side of the ligation junction. For expression profiling, probes are distributed across all isoforms of a gene.

Orientation

The two sides of the ligation junction are called RBD5 (RNA binding domain) and RBD3, respectively. They are named for their orientation relative to the direction of transcription.

Figure 2. (A) Padlock probe contains reverse complementary sequencings to bind to RNA target site. (B) Probe sequence orientation contains the RBD5 and RBD3 on both sides of the ligation junction.

The panel designer accepts the sequences of the mRNA target (**transcription orientation**) and reverse complements them to build the padlock probe (Fig 2).

Ligation Junction

The ligase used by Xenium has a preference for specific sequences at the ligation junctions. Below is a table of empirically derived junction preferences, binned into three groups:

Junction Preferences		
Preferred	**Neutral**	**Non-Preferred**
AT, TA, GA, AG	TT, CT, CA, TC, AC, CC, TG, AA	CG, GT, GG, GC

The probe sequences provided should have one of these dimers at the ligation junction of the sequences provided (the last base of RBD5 and the first base of RBD3; Fig 2B). These junctions are provided in relation to the template strand (transcription direction of the RNA sequence).

While it is best to use the preferred junctions, not all designs are possible with them. It is strongly recommended to avoid the nonpreferred junctions.

Melting Point

It is recommended that the overall Tm of the full-length sequence is between 68°C and 82°C, or between 50°C and 70°C for each arm separately, as calculated by primer3 (https://primer3.org). The web-based version of primer3 does not accept sequences over 36 bp, but the python package `primer3-py` does. There are no specific GC content requirements, but those that fall outside the Tm parameters should be avoided.

The specific parameters used in primer3 are shown below. These parameters can be executed by calling `primer3.calc_tm(probe_seq, **PRIMER3_CALCTM_ARGS)`.

```
PRIMER3_CALCTM_ARGS = {
        "mv_conc": 10,
        "dv_conc": 20,
        "dntp_conc": 10,
        "dna_conc": 5000,
        "salt_corrections_method":
"schildkraut"
}
```

Probe Specificity Evaluation

Probe designs should be evaluated for specificity. If significant homology to a different part of the transcriptome is present, ensure that there are at least four mismatches combined across both arms of the probe. Mismatches closer to the ligation junction disfavor nonspecific ligation. Indels are better than mismatches. Use BLAST or a similar alignment tool to check for off-target homology. When checking for off-target homology, it is important to align to the transcriptome, not the genome.

Other Guidance for Probe Design

- Common rules for oligonucleotide design should be followed, such as avoiding homopolymer repeats or sequences with low complexity.
- Target the coding sequence of a gene instead of the untranslated region if possible.
- Avoid five or more consecutive G bases.
- If multiple probes are being designed for the same target, they should be spaced at least 10 bp apart. For reference, Xenium pre-designed panels default to eight probe sets per gene for the best sensitivity.

Input to the Panel Designer

After probe designs have been finalized, construct a CSV file containing the sequences. This CSV file has three required columns (and a fourth optional), with the following headers:

1. **name**: Equivalent to a gene name. Defines the unique ID that these sequences will appear as in the panel design web summary.
2. **rbd5**: RBD5 sequence from above (Fig 2).
3. **rbd3**: RBD3 sequence from above (Fig 2).
4. **expression**: An optional field that tells the panel designer how highly expressed the targeted RNA is expected to be. It can be any of: Lowest, Low, Medium, High, Highest; these correspond to quantiles of 25%, 50%, 90%, 99%, and 100%. This field is important to use if the target RNA is expected to be very highly expressed.

Notes about Expression Estimation

If a probe is designed against a very highly expressed gene (i.e., a ubiquitously expressed exogenous sequence), it could impact results across the entire panel, not just that probe. The panel designer uses the expression level of targets on the panel to choose barcodes that make the best use of the optical budget. For targets without expression information, an expression estimate can be made. This expression estimate is global across all cell types, and is quantified relative to every gene in the provided reference. If expression information is available on a cell type basis for the target, it can be included in the reference. Both Gene Name and Gene ID fields must match name field in the CSV.

Next Steps

When the CSV file is complete, contact your Sales Executive. Your Sales Executive will generate a custom panel Design ID. Every custom gene panel has a unique Design ID that is used to track your panel through the design process. If you do not know your Sales Executive, please contact customerservice@10xgenomics.com. Submit a PO for an order containing this Design ID prior to working with our Applied Bioinformatics team to generate your design.

Document Revision Summary

Document Number	CG000683
Title	Species Standalone Custom and Advanced Custom Panel Design for In Situ Xenium v1 Panels Technical Note
Revision	Rev C to Rev D
Revision Date	May 2024

Specific Changes:

- Specify that design details apply to Xenium v1 assay chemistry. Updated Technical Note title to reflect version.
- Added Figure 2B to illustrate probe sequence design.
- Added parameters to use for primer3.

General Changes:

- Updated assay chemistry support and probe design information.

Contact:
support@10xgenomics.com
10x Genomics
6230 Stoneridge Mall Road
Pleasanton, CA 94588 USA